Abstract of the Disclosure

The method combines volume growing with change detection. After an input video is filtered to remove noise, a spatio-temporal data structure is formed from the video frames, and markers are selected. From the markers, volumes are grown using a color similarity based centroid linkage method. Change detection masks are then extracted from adjacent frames in the video using local color features. The change detection masks are intersected with each volume, to determine the number of changed pixels only in portions of the masks that lie within that volume. If the number of changed pixels in the intersection exceeds a threshold, then the volume is identified as a moving object.